

ABSTRACT OF THE DISCLOSURE

Methods for processing speech data are described herein. In one aspect of the invention, an exemplary method includes receiving a text sentence comprising a plurality of words, each of the plurality of words having a part of speech (POS) tag, generating a POS sequence based on the POS tag of each of the plurality of words, detecting a prosodic phrase break through a recurrent neural network (RNN), based on the POS sequence, and generating a prosodic phrases boundary based on the prosodic phrase break. Other methods and apparatuses are also described.